DOCUMENT RESUME

ED 061 560 CG 007 166

AUTHOR Durstine, Richard M.
TITLE Datafiles for Computerized Vocational Guidance: Requirements, Preparation, Use. Information System for Vocational Decisions. Project Report No. 15.
INSTITUTION Harvard Univ., Cambridge, Mass. Graduate School of Education.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
BUREAU NO BR-6-1819
PUB DATE May 68
GRANT OEG-1-6-061819-2240
NOTE 37p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Computer Oriented Programs; Computer Storage Devices; Data Processing; Decision Making; Educational Technology; Guidance; *Guidance Counseling; Information Storage; *Information Systems; *Man Machine Systems; *Occupational Choice; *Occupational Information; Secondary School Students; Vocational Counseling; Vocational Development
IDENTIFIERS Information System for Vocational Decisions; ISVD

ABSTRACT

The Information System for Vocational Decision (ISVD) approach to computer-managed information is explained. Drawing mainly from the experience of preparing datafiles of occupational and military information for the ISVD, the author supports and elaborates on a number of general statements: (1) data should be collected and presented specifically to aid students in their vocational decisions; (2) data should be treated in a systematic and structured form that exploits high speed commutation; (3) ready-made sources of information should be relied upon as far as possible; (4) data should be articulated from distinct and diverse sources into a working whole; (5) datafiles and means of access to them should be prepared separate from one another so they can be used in a variety of combinations; and (6) information given by the system should be suggestive, not prescriptive. Other specific criteria and general rules for their implementation are discussed in the final section of the paper. (TL)

ED 061560

CG
BR6-1819

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

INFORMATION SYSTEM FOR VOCATIONAL DECISIONS

Project Report No. 15

DATAFILES FOR COMPUTERIZED VOCATIONAL GUIDANCE: REQUIREMENTS, PREPARATION, USE

Richard M. Durstine

CG 007 166

CG 007 166

This paper was supported in part by Grant No. OEG-1-6-061819-2240 of the United States Office of Education under terms of the Vocational Education Act of 1963.

Graduate School of Education
Harvard University

May, 1968

DATAFILES FOR COMPUTERIZED VOCATIONAL GUIDANCE: REQUIREMENTS, PREPARATION, USE

Richard M. Durstine
Center for Studies in Education and Development
Graduate School of Education
Harvard University

Introduction and Overview

The Information System for Vocational Decisions (ISVD) project has since its beginning embodied the concept of large, orderly collections of factual information as an important part of its resources. This has led to the acquisition of several such collections of information (datafiles). It has also led to some general understandings about collection, storage, processing, and use of information for computerized guidance. These have now been developed and tested to the extent that a unified discussion of them is possible.

I shall present and explain in this paper the approach to computer-managed information that has thus been arrived at, treating theory and related action, both past and planned.

Throughout I seek consistency with related theory and activities of the ISVD. Innocence of guidance and of computer technology may cause some errors in these areas, though I hope not to the extent of invalidating my major points. These latter are summarized below.

1. Data should be collected and presented specifically to aid students in their vocational decisions.
2. Data should be treated in a form that exploits high speed computation. It hence needs to be systematic and highly structured.

3. The power and flexibility of operation thus gained is worth the price of rigidity of structure.
4. For the present, accuracy is of importance secondary to that of operational design.
5. Ready-made sources of information should be relied upon as far as possible. Direct collection of information is costly.
6. It is important (and possible) to articulate data from distinct and diverse sources into a working whole.
7. A primitive form of mediation between "facts/data" and "information" should be included with the datafiles.
8. Datafiles and means of access to them should be prepared separate from one another so they can be used in a variety of combinations.
9. The information given by the system should be suggestive, not prescriptive. It is the inquirer's responsibility to know and make use of this fact.

Some more specific criteria and general rules for their implementation will be found in the final section of this paper. The intervening discussion supports and elaborates on these general statements, drawing support mainly from the experience of preparing datafiles of occupational and military information for the ISVD.

The Fundamental Task in Preparation of Datafiles

The intended role of factual information in the ISVD implies two special requirements:

1. It should be collected, analyzed, and presented specifically to aid a student ("inquirer") in his vocational decisions.
2. It should be treated in a form that exploits the resources of high speed computation. It needs hence to be systematic and highly structured.

Each of these points is perhaps more significant than it looks.

Much factual information of possible value in personal decisions, particularly occupational decisions, is not generally found in useful form. It tends to be better suited for economists, planners, and employers. It needs special interpretation for profitable use by individuals. The "psychology of occupations" as it has developed is a step in this direction. Also needed is modification of highly detailed information to a form that serves the individual. This is the direction of the work described here.

Second, the need to deal with very large quantities of information, concurrently in many combinations and for many purposes, imposes strict requirements on treatment of that information. Also, a working vocational information system should accept with minimum disruption changes or additions to the information it treats, or to its use. Both these needs lead to involvement of the computer and demand a highly structured system. This degree of structure is the price of size and flexibility. I conjecture that the benefit is often worth that price. It is certainly a possibility worth exploring.

The structure thus imposed on factual information within the ISVD has led to the objection that the result may be "too mechanical." But this is functional and necessary to the approach we are taking. To cover it up

would be troublesome and misleading. To eliminate it would for a long time to come be inefficient and costly. To be openly mechanical in this is a simple matter of honesty. A computerized information system that pretends to be fully human has no more self-evident merit than an airplane with flapping wings or a telephone with moveable lips at the earpiece.

It should be noted for clarity that here "information" generally refers to the "facts/data" of Tiedeman. In its ultimate use it may approximate Tiedeman's concept of "information," but since this distinction is not central to the present discussion, it will not be used, except where Tiedeman's work is specifically referred to. "Information" will thus usually be used here in an unspecific and pedestrian sense, more or less interchangeably with "data."

A Datafile and How it Grew

The development of datafiles within the ISVD began with preparation and use of the project's first datafile in the academic year 1966-67. A brief description of this file will motivate some of the conclusions that have been subsequently reached concerning datafiles in general.

The first datafile will be referred to here as the "850 Titles File." It contains information under fifty-six categories about some 850 occupational titles. The categories are shown in Appendix 1. This title-category structure is a useful--and obvious--one for datafiles in general. To render this concept explicit, an illustrative "mini-datafile" is shown in Figure 1.

TITLE	D.O.T. NUMBER	CODED DATA	DESCRIPTIVE INFORMATION
FOREST ENGINEER	005.187	45 0 5 1	DESIGNS AND OVERSEES CONSTRUCTION OF FACILITIES FOR LOGGING
EGG BREAKER	521.887	4 I 2 9	SEPARATES EGGS FOR USE IN FOOD PRODUCTS

Figure 1

Illustrative Short Datafile Based on Occupational Titles

This tiny illustrative file gives a brief verbal description of the activities of workers in each of two occupations. In addition, in coded form, are given:

1. Dictionary of Occupational Titles (D.O.T.) number of the occupation;
2. Working conditions (first two columns under "Coded Data");
3. Education required (third column under "Coded Data");
4. Seasonality of work (fourth column under "Coded Data").

For example, the code "45" indicates that a forest engineer must use his hands, and be able to speak and hear to do his work. An egg breaker, on the other hand, also works with his hands, but need not speak or hear. This information is contained in the position and identity of the letters and numbers in the file. Hence great care is required in designing the datafile to transmit precisely the intended meaning.

The illustrative datafile also tells that a forest engineer works out of doors (an "O" in an assigned position of the file carries this information), while an egg breaker works indoors (indicated by an "I" in like position). Likewise a "5" and a "2," appropriately placed,

indicate the levels of general education required for each (some college for the forest engineer, and some high school for the egg breaker). Finally, in the last column of code, a "1" indicates some seasonality in the engineer's work. "9" shows no information on this for the egg breaker.

This trivial illustrative file exhibits all the characteristics and uses associated with any datafile of this form. Though the 850 Titles File requires sixteen punched cards for the coded and other information about each title, it and the illustrative file are identical in their essential characteristics.

The mechanics of construction of a datafile in this form are simple but strict. The way in which each piece of information is to be expressed must be unambigously fixed. Words, numbers, or code may be used. In each case, exact meaning must be decided upon and stated. Words have the most flexibility, of course, and codes the least. Codes are the most economical of space. Since the structure of the datafile is highly rigid, changes or deletions can be made readily. Likewise, additions of titles or of categories are conceptually simple and mechanically straightforward. This flexibility of modification is one advantage bought with rigidity of form. This is of great value for any datafile that can be expected to change in time.

Another point, perhaps less obvious and surely more controversial, is that accuracy of information is of secondary importance for the time being. It is not unimportant, but less so than structure and the means to use it. For this reason, though a diligent attempt was made at accuracy in preparation of the 850 Titles File, no great effort is presently

given to its modification and updating. This is a mechanical task, of less immediate interest than learning to use the file in an imaginative and flexible way.

Preparation of a large datafile of factual information reveals the dependence of such files on large blocks of information prepared for other purposes. The work of agencies such as the Bureau of Employment Security and the Bureau of Labor Statistics is invaluable because many of their results can be taken readily and inexpensively into the file. This suggests three further criteria of datafile preparation.

1. Direct preparation of information in large quantities is costly, so that ready-made information sources should be relied upon as far as possible. Some information will eventually have to be gathered and prepared explicitly for the working ISVD, but this must be selected with great care in terms of its cost and its usefulness.
2. Since ready-made information will seldom exactly suit the purposes it is to serve (unless they are the purposes for which it was explicitly prepared), it must be modified to a form as appropriate to the new use as possible.
3. The need to use as many information sources as possible makes necessary the ability to articulate diverse sources or collections of information into a working whole.

The above three requirements arise directly from the need to make maximum use of available resources. This need might be classed by some as regrettable. It is, however, so universal and unavoidable that it is merely one more fact in preparing information for use by the ISVD.

given to its modification and updating. This is a mechanical task, of less immediate interest than learning to use the file in an imaginative and flexible way.

Preparation of a large datafile of factual information reveals the dependence of such files on large blocks of information prepared for other purposes. The work of agencies such as the Bureau of Employment Security and the Bureau of Labor Statistics is invaluable because many of their results can be taken readily and inexpensively into the file. This suggests three further criteria of datafile preparation.

1. Direct preparation of information in large quantities is costly, so that ready-made information sources should be relied upon as far as possible. Some information will eventually have to be gathered and prepared explicitly for the working ISVD, but this must be selected with great care in terms of its cost and its usefulness.
2. Since ready-made information will seldom exactly suit the purposes it is to serve (unless they are the purposes for which it was explicitly prepared), it must be modified to a form as appropriate to the new use as possible.
3. The need to use as many information sources as possible makes necessary the ability to articulate diverse sources or collections of information into a working whole.

The above three requirements arise directly from the need to make maximum use of available resources. This need might be classed by some as regrettable. It is, however, so universal and unavoidable that it is merely one more fact in preparing information for use by the ISVD.

The problem, then, is to handle information in a way that will best serve the uses of the system, given existing technological and economic limitations. In the following pages some steps toward a working solution will be presented.

Clearly the comments here are not limited to information about occupations, though they are derived therefrom and are hence particularly suited to that application. Any data that can be readily described in terms of titles and categories can be treated in this way. Thus these comments have considerable generality. The extent of their applicability must be judged in each individual case of course.

Further issues of collection and storage of information will be discussed later. First, however, it will be helpful to review some theoretical characteristics of the ISVD as they relate to the preparation of datafiles.

Role of Datafiles in the ISVD

In a recent paper ("The Role of Decision-Making in Information Generation," ISVD Project Report 12), Tiedeman has given "primary data files" a central role in his operational definition of the ISVD working system. These primary data files are the datafiles referred to in the present paper, of which the file of occupational titles described above is an example. The centrality of these datafiles within the ISVD thus requires their careful articulation with the system as a whole. The considerations that thus arise are of five kinds.

First, the datafiles (primary data files, in Tiedeman's terminology) will function in the system together with some form of mediation to help convert the "facts/data" of the files to "information" by individual inquirers in the system. A major task of the system is to provide such

mediation. Although the preparation of datafiles thus does not require the preparation of such mediating elements, the process of mediation will be helped if the datafiles are properly designed, and if some primitive form of mediation is provided with them. This both justifies and motivates the first steps that were taken toward use of the datafiles described above. These were:

1. Presentation of the contained information in English language form;
2. Provision for selection of information in terms of certain of its more important characteristics.

These two steps toward conversion of "facts/data" to "information" will be elaborated later.

Second, the process of decision-making by the individual inquirer is conceived in the design of the system as taking place at a number of discrete "discontinuities." A central principle of the ISVD is that the inquirer should be helped to maximize his awareness of and participation in these discontinuities. The datafiles in use at any time will be identified in part by the discontinuity the inquirer is dealing with. As the system becomes adept at dealing with a variety of discontinuities, the structure and use of the datafiles to serve this end must be more and more highly refined. Thus the datafiles must remain flexible to varying demands depending on which discontinuity is being served. This is further reason why the highly structured form chosen for the occupational datafile is appropriate to economical satisfaction of the demands of the system. It permits a single datafile to serve a variety of discontinuities.

Parenthetical note: While it is necessary to speak briefly here of the theory underlying the ISVD, this is not meant to be a full or definitive statement of that theory. Other documents should be referred to for that purpose. The intent here is to suggest that the decisions made with regard to datafiles have been appropriate to the overall needs and aims of the system.

Third, the theory behind the ISVD prescribes that the inquirer be encouraged and taught to deal with his discontinuities in terms of a paradigm of several sequential stages. Again the datafiles will be called upon to perform differently depending on which of these stages the inquirer is in. Since at any moment each inquirer is likely to be dealing with a different discontinuity and to be at a different stage of the paradigm, great flexibility is demanded of the datafiles. The structure of datafiles thus far provided is a step in meeting these needs.

Fourth, the capacity of the inquirer to absorb and use information should have some influence on which information is provided to him, and how. The same information might be presented in a number of ways, depending on the needs and capabilities of the inquirer. If the datafile itself and the means of getting at it are separate (i.e., if more than one means of access and presentation can be adjoined to a single datafile) flexibility in use of a single datafile will be increased.

Fifth and finally, several means for mediation of "facts/data" to "information" will be used by the system. These include direct teaching of concepts, simulation, and real experience with decision-making. The datafiles should stand ready to serve these various pedagogical modes. This reinforces the requirements of flexibility, and of access to the files separate and distinct from the files themselves.

A Young Datafile's First Steps

It is clear, then, that the concept of the ISVD and of the operation of datafiles within it are both highly ambitious. Considerable time will be needed to bring them to fruition. It is hardly possible to conceive of there ever being a final system, but rather of a working and evolving system. It is important to distinguish dreams from forseeable accomplishments, and both from present achievements. The latter are important in that they help to confirm and give hope to the dreams. They also point the way to activities needed to bring forseeable accomplishments to reality. It is thus desirable to consider applications of the 850 Titles File as it has developed, not as a final working thing, but as a first step, and as a suggestion of best directions for future development.

It will be recalled that the 850 Titles File consists of some 850 titles with information in fifty-six categories (prepared for computer use by placing the data for each title on sixteen punched cards; transferred subsequently to magnetic tape). It will also be recalled that this information is almost entirely in compact coded form of little direct use to an inquirer of the ISVD.

An obvious question in converting this datafile to use by inquirers is how to present its contents in a form suitable for human comprehension. The answer chosen is fairly simple. The mechanical problem of presentation of coded information is that of converting it to English or some other readable form. A second problem is that of choosing and interpreting information in a way that is maximally meaningful and minimally misleading. The mechanical problem is simpler, and will be discussed first and more fully.

The form adopted for presentation of information from the 850 Titles File can be readily described in terms of the mini-datafile presented earlier. For each of two occupational titles, this file contains a brief description of the occupation plus certain coded information. For the moment it is not important whether this information is either appropriate or accurate. The mini-datafile can answer questions such as "What does an egg breaker do"? "How much education does a forest engineer need"? "Where does an egg breaker work"? Answering such questions requires three steps:

1. Identify which category(ies) of stored information contains the answer to the stated question.
2. Find the appropriate information by title and category.
3. Present answer in terms understandable to the inquirer.

The first two of the above steps are technical ones, to be taken through accurate identification and location of the stored information. The last can be made by constructing a suitable sentence. An example will suffice. To answer "Where does an egg breaker work"?, the system must first identify where information of this type is kept, if it is kept at all. In the illustrative mini-datafile, this is the third column of code, in which an "I" or an "O" is found. From this it is possible to respond "An egg breaker works indoors."

More generally, it can be said: "A (Title) works (x.)" This answers the question "Where does a (Title) work"? The system will place "indoors" in the sentence if the datafile contains an "I" in the appropriate location, and with "outdoors" if there is an "O." If there are further possible situations, such as "both," such a code must be defined, and words

provided to interpret that code in the sentence. A special case is the possibility of a blank, or of an illegal answer, in which case some sort of null response would be called for.

The above procedure can be applied with considerable generality, given three things:

1. A skeleton sentence (e.g., "A (Title) works (x).") which makes sense for each possible insertion of (Title) and (x).
2. An English interpretation of the content of the datafile for each possible content. This specifies "x" in the above sentence.
3. A substitute message if the coded information is unavailable or inappropriate.

Satisfaction of the above requirements calls for some precision and care, but is by no means impossible or even particularly difficult. Once the rules are set, new titles can be added--or information added, deleted, or changed in the file. Likewise, the form of presentation can be altered without changing the coded information. Thus to a great extent the content and use of the datafile can be separated, with resultant valuable flexibility.

It is now a direct step to answering the request "Tell me about the occupation egg breaker" or "Tell me about the occupation forest engineer." In either case the answer would be built of the various sentences that answer individual questions about the named occupation. The result will of course seem somewhat mechanical, but with care in preparation it should not be unbearably so. Descriptions for the 850 Titles File have been prepared in this form.

The second question, that of choice and interpretation of information, must be resolved over a longer term, and is much more difficult to handle adequately and honestly. The approach taken so far has been to use whatever data is available, and to be forthright about what it does and does not say. The flexibility designed into the datafile and into the presentation of its contents will then facilitate exploitation of improvements in quality of information as they become available.

The structure of the 850 Titles File, its contents, and the presentation of its information in the form described above, have been described in detail in working documents of the ISVD project. Further elaboration of these topics will not be given in this review article. The question of access to the information in this file will similarly be treated in a brief manner.

Questions about any individual job readily come to mind, and can be listed more or less briefly, particularly given knowledge of the limited scope of information on each title. The choice of occupation(s) to ask about is far less obvious. To simply present a list of 850 titles and say "You may ask about any of these," would be folly. Such a list in its entirety is of little use. Some method of selective access is needed.

The most direct way of making this choice is in terms of selected characteristics of the titles in the datafile. Again I resort to the mini-datafile for illustration. The inquirer might ask "Which occupations involve indoor work and require less than a high school education"? The answer would be, "egg breaker." Let us examine this procedure and its limitations.

1. Response to the question can only be in terms of the occupational titles on the list, which is a limitation of the occupational titles approach. One way of relaxing this limitation will be dealt with presently.
2. Second is the matter of which questions the inquirer can ask the datafile. At the present stage of use of the 850 Titles File, questioning is in terms of responses to a set of multiple choice questions. In a fully free situation, he might make other requests, to which the datafile could not respond. In that case a "don't know" or "unfortunately an answer on this subject is not available" message would have to be prepared.
3. A third eventuality is exhaustion of the file. With the mini-datafile, there are, for example, no outdoor jobs that require less than a high school diploma. With a suitable message, this ceases to be a problem.

Search of the file by the computer for suitable titles is a straightforward matter, subject to the limitations mentioned above, and will see realization in the first working prototype of the ISVD.

It is important to note throughout that the attitude in presentation of data by the system is not prescriptive, but suggestive. The inquirer is not told that he must pursue life as an egg breaker, or even that this is desirable. He is merely informed of this alternative, and that it meets his apparent occupational specifications. A description of the occupation, prepared by the system as described above, tells him other facts about

egg breakers (or forest engineers, or whatever), and gives him references to further information. He is then free to try out other sets of specifications, or to accept or reject further consideration of any title, either before or after he has received a description of it.

It is important for the inquirer to realize that the system tells what it is asked, and that it can tell no more than it knows. It is his responsibility to understand and make use of this fact. If handled properly, a system with incomplete information can be of use to the inquirer. This understanding is imperative, because full information will never be impossible. A large and growing body of information will be a reality, however, and can be made responsive to the inquirer's needs, if he knows how to deal with it properly.

Extension to Other Datafiles

Design for an evolving structure of datafiles within the ISVD can be based on experience with the 850 Titles File. First, however, a final dimension must be added to the problem, namely the possibility of a complex datafile composed of two or more files of the type already described.

This extension will be introduced by an example from the actual development of datafiles for the ISVD. The initial collection of information was made in terms of the 850 Titles File. Information from a variety of sources was coded into a highly structured framework as already described. At subsequent times other information also became available, including:

1. The supplementary volume of the Dictionary of Occupational Titles (D.O.T.);
2. Classification by Anne Roe's categories and levels of some 800 occupational titles;

3. Forecasts of demand to 1975 by occupation and industry of some 160 occupational groups and some 120 industrial groups.

The availability of this new information posed an important problem in information handling. In each new case the titles covered were substantially different from one another and from those in the originally prepared 850 Titles File. In the case of the supplement to the D.O.T., virtually all jobs of the earlier datafile were included, so the additional information could be adjoined readily to the 850 Titles File. Waste arises from the fact that this leaves more than 90 percent of the new information unused.

The information about Roe categories and levels (item 2 above) covers about the same number of titles as the original datafile. The two lists are far from being equivalent, however. The bulk of the new information can be put to use by including it where appropriate, and leaving Roe category unspecified elsewhere.

In the case of the forecast information (item 3 above), the situation is considerably different. This information is given not by occupational title, but by occupational group. The titles used for groups of occupations and of industries exhaust all possibilities, referring sometimes to titles as "not elsewhere classified." Whereas there exist occupations that are not found in the D.O.T., there are in principal none that do not fall into one of this smaller set of occupational groups. The two classification systems are thus qualitatively different, and must clearly be treated separately.

Information by occupational title and information by occupational group can be articulated by treating one as an example of the other. Thus

an example within the group "structural metal workers" is "pneumatic riveter," which is a specific title in the 850 Titles File. Likewise "pneumatic riveter" can be identified as a member of the group "structural metal workers," and through that identification other example titles can be found. Thus the disparity between the two lists is turned to good use. It adds flexibility to the search for titles of interest.

The above experience can be summarized in general terms as follows.

1. A datafile might very well consist of two or more pieces (title-category blocks) that have distinctly different titles and categories of information included in them.
2. Translation between these can be facilitated by suggesting the titles of each block that correspond to each title of the other. These translations need not be unique in either direction.
3. Free use and flexibility of the datafiles can and should be encouraged by making these translations suggestive rather than prescriptive.

A second case of articulation of the sort described above has been carried out between the occupational datafile and the military datafile. The latter lists categories of assignments for enlisted men. This articulation promises to be helpful to the use of both these files.

Specifications for the Continuing Development of Datafiles

Datafiles for the ISVD and procedures for their use will undergo continuing development. There is no forseeable end to additions or alterations to this material and its uses. Any closed system of datafiles is to

be shunned. A general plan for development is needed in which the datafiles are continuously operative, but in which new information can be accepted readily and with minimum disruption of operation. Such a plan is suggested by the experience of datafile development described above. It derives explicitly from the datafiles on occupations and military service, but should apply readily to those on education, family living, and other topics an information system like the ISVD might eventually encompass. Needed are an ability to arrange the information in terms of titles and categories, as mentioned earlier; and to translate among the various sets of titles. The files thus can be readily articulated among themselves. They potentially, therefore, are not a set of datafiles, but one large complex file.

Experience so far with datafile development suggests certain criteria and a set of rules to meet these criteria. These may seem self evident or trivial in retrospect. But they were by no means obvious during the development of the datafiles described here. They might also seem excessively general in form, but this is intentional, to allow application to a wide range of cases.

1. Datafiles should adapt to different use depending on which discontinuity and stage of decision-making they are serving for each individual inquirer.
2. Means of entry to each datafile should be independent from the file itself, so that either the means of entry or the file itself can be changed without disturbing the other.

3. Likewise, form of presentation of information from a datafile should be independent from the file itself for the same reasons.
4. Files should be designed so that additions and updating are possible without undue disruption of the existing files or of their operation.

These criteria, recognized as important in development of the existing datafiles on occupations and military service, must be interpreted individually for each datafile. The following rules have been helpful in this regard.

1. Collect information by blocks in which information in well defined categories is given for a set of well defined titles.
2. Entry to these blocks can be according to selected categories, or by title.
3. Coded information should be presented to the inquirer in English or some other readily understandable form. This in general will require structured formats within which to present the information.
4. Free access between separate blocks of information should be facilitated through explicit translation from the titles of one to the titles of the other. Properly done, this will provide flexibility and freedom in finding and taking information from the datafiles.

Much work remains in the preparation of datafiles. The job will probably never be over, since updating and additions can be expected to

go on so long as there is a system. Three immediate needs, which set the stage for the near future, follow.

1. More accurate information, when and as this becomes available in readily usable form.
2. More appropriate information, aimed at individual decision-making rather than economic or large scale planning.
3. Less structured entry to the information. This is mainly a matter of information processing capability and is beyond the scope of this paper.

Appendix 1

Categories of Information from The 850 Titles File
and from the D.O.T. Supplement

Category	In 850 Titles File	In D.O.T. Supplement
1. Occupational title	yes	yes
2. Alternate titles	up to five	
3. Entry occupations	up to four	
4. Higher occupations	up to four	
5. Industries where found	up to four	one only
6. Brief verbal description	yes	
7. Worker trait sector (from D.O.T.)	yes	yes
8. Aptitudes: General intelligence	five levels	five levels
9. Aptitudes: Verbal ability	five levels	five levels
10. Aptitudes: Numerical ability	five levels	five levels
11. Aptitudes: Spatial perception	five levels	five levels
12. Aptitudes: Form perception	five levels	five levels
13. Aptitudes: Clerical perception	five levels	five levels
14. Aptitudes: Motor coordination	five levels	five levels
15. Aptitudes: Finger dexterity	five levels	five levels
16. Aptitudes: Manual dexterity	five levels	five levels
17. Aptitudes: Eye-Hand-Foot coordination	five levels	five levels
18. Aptitudes: Color discrimination	five levels	five levels
19. Interest Preferences	up to three	
20. Occupational situations	up to three	up to five
21. Strength required	five levels	five levels

Appendix 1 (continued)

Category	In 850 Titles File	In D.O.T. Supplement
22. Physical demands	up to five	up to five
23. Required high school courses	36 possibilities	
24. Salaries and wages	three levels	
25. Required academic ability	four levels	
26. Required mechanical ability	four levels	
27. Required social intelligence	four levels	
28. Required clerical ability	four levels	
29. Required musical ability	four levels	
30. Required artistic ability	four levels	
31. Required physical ability	four levels	
32. Most common prior activity	up to two	
33. On-the-job training	eight levels	See item 63
34. Formal education required	eight levels	See item 62
35. Formal education preferred	eight levels	See item 62
36. Recommended School courses	up to four	
37. Short training courses	up to two	
38. Distribution of sexes	five levels	
39. Minimum age	yes	
40. License requirement	yes	
41. Union membership	yes	
42. Working conditions	up to six	up to six
43. Place of work (indoor, outdoor)	yes	yes
44. Weekend work	three levels	

Appendix 1 (continued)

Category	In 850 Titles File	In D.O.T. Supplement
45. Hours of work	three levels	
46. Requirements for travel	three levels	
47. Seasonality	yes	
48. Basis of income	four levels	
49. Incentives and fringe benefits	up to four	
50. Opportunities for promotion	four levels	
51. Paths for promotion	up to two	
52. Opportunities for mobility	four levels	
53. Possibility of self-employment	yes	
54. Growth rate of occupation	four levels	
55. Demand for workers	five levels	
56. Trend of wages	three levels	
57. Worker functions--data		ten levels
58. Worker functions--people		nine levels
59. Worker functions--things		eleven levels
60. Work fields		up to two
61. D.O.T. number	yes	old and new
62. General educational development	See items 34,35	six levels
63. Specific vocational preparation	See item 33	nine levels
64. Materials, products, subject matter		up to two

Appendix 1 (continued)

NOTES:

There is some redundancy among the categories of the 850 Titles File that must be worked out in its use.

The two data files described here are merged into one for operation within the ISVD.

In all cases there is a possibility that "not applicable," "no information" or some similar null message can be coded as appropriate.

Appendix 2

Relation of Occupational Titles of The 850 Titles File
to Occupational Groups of The Forecasting File

The three digit coding of occupational groups shown below is used for forecasting and for certain summary statistics. The number of titles from the 850 Titles File that fall into each group is given in the right hand column.

Occupational Group Code Number	Group Title	Number of Titles in 850 File
100	Other Professional & Technical Workers	17
101	Accountants and Auditors	3
102	Airplane Pilots and Navigators	2
103	Architects	3
104	Workers and Teachers in Arts and Entertainment	15
105	Clergymen	4
106	Designers (ex. Design Draftsmen)	5
107	Editors and Reporters	9
108	Lawyers and Judges	2
109	Librarians	4
111	Personnel and Labor Relations Workers	4
112	Photographers	2
113	Social and Welfare Workers	2
120	Other Engineers	9
121	Aeronautical Engineers	1

Appendix 2 (continued)

Code Number	Group Title	Number of Titles in 850 File
122	Chemical Engineers	1
123	Civil Engineers	3
124	Electrical Engineers	6
125	Industrial Engineers	4
126	Mechanical Engineers	3
127	Metallurgical Engineers	2
128	Mining Engineers	1
130	Other Natural Scientists	3
131	Chemists	4
132	Agricultural Scientists	7
133	Biological Scientists	12
134	Geologists and Geophysicists	3
135	Mathematicians	2
136	Physicists	1
140	Other Technicians	3
141	Draftsmen	1
142	Surveyors	3
143	Air Traffic Controllers	2
144	Radio Operators	3
150	Other Medical and Health Workers	10
151	Dentists	1
152	Dietitians and Nutritionists	2
153	Professional Nurses	1

Appendix 2 (continued)

Code Number	Group Title	Number of Titles in 850 File
155	Optometrists	1
156	Osteopaths	1
157	Pharmacists	1
158	Physicians and Surgeons	2
159	Psychologists	3
161	Medical and Dental Technicians	5
162	Veterinarians	1
170	Other Teachers	4
171	Elementary Teachers	1
172	Secondary Teachers	1
173	College Teachers	2
180	Other Social Scientists	5
181	Economists	2
182	Statisticians and Actuaries	4
200	Other Managers, Officials and Proprietors	40
201	Railroad Conductors	2
202	Creditmen	2
203	Ships Officers, Pilots, and Engineers	6
204	Purchasing Agents	3
205	Postmasters and Assistants	1
300	Other Clerical Workers	26
301	Accounting Clerks	3

Appendix 2 (continued)

Code Number	Group Title	Number of Titles in 850 File
302	Hand Bookkeepers	2
303	Bank Tellers	1
304	Cashiers	2
305	Mail Carriers	1
306	Postal Clerks	2
307	Shipping and Receiving Clerks	2
308	Telephone Operators	3
310	Stenographers, Secretaries, and Typists	6
320	Office Machine Operators	15
400	Sales Workers	29
500	Other Craftsmen and Foremen	91
501	Bakers	3
502	Cabinetmakers	1
503	Cranemen, Derrickmen, and Hoistmen	6
504	Glaziers	1
505	Jewelers and Watchmakers	8
506	Loom Fixers	2
508	Opticians and Lens Grinders	5
511	Log and Lumber Inspectors	1
512	Other Inspectors	1
513	Upholsterers	3
520	Other Construction Craftsmen	9
521	Carpenters	2

Appendix 2 (continued)

Code Number	Group Title	Number of Titles in 850 File
522	Brickmasons, Stone and Tile Setters	4
523	Cement and Concrete Finishers	3
524	Electricians	2
525	Excavating & Grading Machine Operators	8
526	Painters and Paperhangers	2
527	Plasterers	2
528	Plumbers and Pipefitters	3
529	Roofers and Slaters	1
531	Structural Metal Workers	6
540	Other Foremen	(None explicitly)
550	Other Metalworking Craftsmen (except Mechanics)	20
551	Skilled Machining Workers	5
552	Blacksmiths, Forgemen, and Hammermen	5
553	Boilermakers	1
554	Heat Treaters, Annealers, Temperers	1
555	Millwrights	1
556	Metal Molders (except Coremakers)	3
557	Wood and Metal Pattermakers	3
558	Rollers and Roll Hands	3
559	Sheet Metal Workers	1
561	Toolmakers, Diemakers, and Setters	2
570	Other Mechanics and Repairmen	32
571	Airplane Mechanics	2
572	Motor Vehicle Mechanics	2

Appendix 2 (continued)

Code Number	Group Title	Number of Titles in 850 File
573	Office Machine Servicemen	1
574	Radio and Television Mechanics	1
575	Railroad and Car Shop Mechanics	1
580	Other Printing Trades Craftsmen	4
581	Compositors and Typesetters	5
582	Electrotypers and Stereotypers	2
583	Engravers (except Photo)	(None)
584	Photoengravers	7
585	Pressmen and Plate Printers	3
590	Other Transportation and Public Utility Craftsmen	10
591	Linemen and Servicemen	10
592	Locomotive Engineers	1
593	Locomotive Firemen	1
600	Other Semi-skilled Workers	113
601	Automobile Parking Attendants	(None)
602	Blasters and Powdermen	1
603	Laundry and Dry Cleaning Operatives	3
604	Mine Operatives and Laborers	1
605	Meatcutters (except Meatpacking)	3
606	Asbestos and Insulation Workers	1
610	Other Transportation and Public Utilities Operatives	4
611	Truck, Bus, and Tractor Drivers	4
612	Deliverymen, Routemen, and Cab Drivers	3

Appendix 2 (continued)

Code Number	Group Title	Number of Titles in 850 File
613	Railroad Brakemen and Switchmen	2
614	Power Station Operators	3
615	Sailors and Deckhands	(None)
620	Other Semiskilled Metalworking Occupations	4
621	Metalworking Assemblers	8
623	Furnacemen, Smeltermen, and Pourers	3
624	Metal Heaters	5
625	Machine Tool Operators	(None)
626	Metalworking Inspectors	1
627	Welders and Flamecutters	3
628	Electroplators	1
629	Electroplater Helpers	(None)
630	Other Semiskilled Textile Occupations	5
631	Knitters, Loopers, and Toppers	7
632	Textile Spinners	1
633	Textile Weavers	1
634	Sewers and Stitchers	2
700	Other Service Workers	10
701	Airline Stewards and Stewardesses	1
702	Hospital and Other Institution Attendants	2
703	Charwomen and Cleaners	(None)
704	Janitors and Sextons	2
705	Practical Nurses	1

Appendix 2 (continued)

Code Number	Group Title	Number of Titles in 850 File
710	Private Household Workers	(None)
720	Other Protective Service Workers	1
721	Firemen and Fire Protection Workers	1
722	Policemen, Detectives, Marshalls and Sheriffs	3
723	Guards, and Watchmen	2
730	Other Food Service Workers	1
731	Bartenders	1
732	Cooks (except Private Household)	2
733	Counter and Fountain Workers	1
734	Waiters and Waitresses	2
800	Laborers (except Farm and Mine)	(None)
900 -	Farmers and Farm Workers	14

Appendix 3

Summary of Datafiles on Occupations and Military Service
Presently in the ISVD

File Identification	Number of Titles (approx.)	Number of Categories	Summary Description
850 Occupational Titles	860	56	See Appendix 1
DOT Supplement	13,800	27	See Appendix 1
Roe Groups	800	2	Roe Category and Level
Forecasting	160	160	High and low estimates of employment by year, 1960-80, for six selected industries (out of 120 possible industries)
Military enlisted	170	varies by service	About ten categories for each service
Military Officers	Not applicable	Not applicable	This information is presented in script form
References	10,000	1	Provides link to library of occupational information at Newton High School

Appendix 3 (continued)

Future development of datafiles in the ISVD can be expected to center around topics such as:

1. Local information and forecasts;
2. Career patterns (e.g., "job ladders");
3. Estimates of future relative demand by skill and type of training;
4. Information about industries;
5. Action steps needed to carry out occupational plans.